DOCUMENT RESUME

ED 401 296 TM 025 839

AUTHOR Mead, James V.; Simon, Stacy
TITLE Developing a Mid-Range Theory To Make Sense of Scaling-Up School Reform: A Cross-Site Case Study of Chicago Public Schools.
SPONS AGENCY Chicago Panel on School Policy, IL.; Lloyd A. Fry Foundation, Chicago, IL.; Spencer Foundation, Chicago, Ill.
PUB DATE Apr 96
NOTE 27p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Achievement Gains; Case Studies; *Educational Change; Educational Theories; Elementary Secondary Education; *Outcomes of Education; Public Schools; *School Restructuring; *Urban Schools
IDENTIFIERS *Chicago Public Schools IL; Fractals; Reform Efforts; *Theory Development

ABSTRACT

This paper reports the results from a cross-case analysis of a series of individual case studies in "successful" Chicago (Illinois) schools. Success was defined as demonstrating learning gains in student assessments. Taking this output approach, researchers confirmed the frequently reported findings from reform process case studies that successful schools create, maintain, and communicate a vision with an educational focus in the "steady work" fashion. Furthermore, these successful school used a process of local adaptation to make reform efforts fit into their environment. These are significant findings in that they replicate findings normally attributed to case studies that examine a singular reform effort or process. It is argued that these findings and backward mapping theory lead to the consideration of the output of reform, rather than its processes, as the appropriate target for future mid-level or system-wide theory development. The paper concludes with some speculation about using fractal modeling techniques to simulate reform output over large-scale systems. (Contains 1 figure and 22 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 401 296

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
JAMES V. MEAD
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Developing a mid-range theory to make sense of scaling-up school reform: A cross-site case study of Chicago Public Schools[1].

James V. Mead & Stacy Simon

James V. Mead
Boston, MA
Stacy Simon
Chicago, IL

[1]This research was sponsored in part by the Spencer Foundation, Lloyd A. Fry Foundation and the Chicago Panel on School Policy. The views expressed in this paper are those of the authors and do not represent the position, policy, or endorsement of the Spencer Foundation, Lloyd A. Fry Foundation or the Board of the Chicago Panel on School Policy.

TM025839

BEST COPY AVAILABLE

Abstract

Developing a mid-range theory to make sense of scaling-up school reform: A cross-site case study of Chicago Public Schools.

This paper reports the results from a cross-case analysis of a series of individual case studies in "successful" Chicago schools. Success in this case was defined as schools that had demonstrated learning gains in their student assessments. Taking this output approach the researchers confirmed the frequently reported findings from reform process case studies that successful schools create, maintain, and communicate successfully a vision with an educational focus in the "steady work" fashion. Furthermore, these successful schools used a process of local adaptation to make reform efforts fit into their environment. These are significant findings in that they replicate findings normally attributed to case studies that examine a singular reform effort or process. The authors argue that these findings and backward mapping theory lead them to consider reform's output not processes as the appropriate target for future mid-level or system-wide theory development. The paper concludes with some speculation about using fractal modeling techniques to simulate reform output over large scale systems.

J.V. Mead & S. Simon.

Developing a mid-range theory to make sense of scaling-up school reform: A cross-site case study of Chicago Public Schools[1].

Introduction

Eyvind Kelda is a man literally caught in the tides of history. He managed to evade the fiery end planned by King Olaf that engulfed his fellow sorcerers and witches. Unfortunately his lucky escape made him supremely confident in his ability to control events in his life. Eyvind added an insult to the king's injury by telling all who would listen that he intended to practice his spells and he could never be caught. It is said the king was greatly vexed and he mounted an expedition to catch the boastful warlock. The king captured Eyvind when his own spells trapped him. We last hear of Eyvind and his friends about to drown as they stand chained to rocks that will be covered by the incoming tide.

"Did his [Eyvind] mind rise above a thirst for vengeance and pity for his companions to grand metaphysical speculation?" asks Eckland (1993, p. 35). I wonder if more detail in Eyvind's story could help us? We would like to decide whether random events were responsible for his impending death or whether a series of policies he pursued led inevitably to the fatal early morning dip in the sea. As is probably true in his case and when we consider the spread of reform in Chicago's successful schools we need a theory that accommodates both deliberate policy and chance. On the one hand, it seems as if a series of random events created the school's success. On the other hand, we could see each case as illustrative of some rules, general principles, or stories that if we only understood them then we could exercise control over our educational fate. The problem was that each time we attempted to scale-up a particular theme or feature to develop a cross-case process model (mid-level theory) the theme lost vital content. Intuitively we came away from the experience feeling that a mid-level reform model should embrace randomness and determinism if we are to make sense of what we see in schools.

Hargreaves (1985) asked for relief from what he described as a void in the policy continuum between micro case studies of reform at the school or classroom level and macro reform knowledge at the federal and state levels. Elmore and Fuhrman (1995) note that "we lack a mid-range theory about the types of inputs [or processes] that matter in student achievement. We lack the knowledge of whether those inputs are amenable to influence by state or local policy and if so what kinds of policy might work." The task is to find a diffusion model that makes sense of how reforms spread between parts of the educational system.

---

[1]This research was sponsored in part by the Spencer Foundation, Lloyd A. Fry Foundation and the Chicago Panel on School Policy. The views expressed in this paper are those of the authors and do not represent the position, policy, or endorsement of the Spencer Foundation, Lloyd A. Fry Foundation or the Board of the Chicago Panel on School Policy.

People interested in school reform hope to achieve systemic restructuring through the application of a mid-level process theory. However, those involved with reform on the state or district scales of school policy have instruments with questionable efficacy (Hall, 1995). To paraphrase Kirst (1995), the route between the state legislature or the district office and the classroom takes many twists and turns with state policy creeping into the classroom. High-level state or federal policies and theories have a long and difficult journey into individual teacher practice. However, at a district, school, or classroom scale, there lies a patchwork of restructuring policies, that enjoy varying degrees of implementation by districts, schools, or individual educators across and within districts. Wang, Haertel, and Walberg's (1995) review of the effective schools literature mentions 228 possible school scale reform influences on student achievement. This gives a sense of the intricacy of the patchwork at the district, school, and classroom scale of operation.

Our individual school cases illustrated this patchwork of restructuring but this confusion gave us an insight into understanding how reform reveals itself at a scale beyond individual schools. The cross-case analysis did not provide a better understanding of a mid-level reform theory process but forced us consider the pattern of products that these reform processes yield across a district or multiple school sites. Mid-level reform theory is not scaling-up to a district level the messy complexities of process that exist at the school and classroom level. Neither do we see a mid-level theory as successfully scaling-down state or national reform policy. A mid-level theory is not an intermediary between reform activities that take place at high and low levels of the educational system. Instead we argue that the appropriate role for a mid-level theory is to model the educational output produced by the interaction of low and high level reform processes as they spread across a district or state.

Processes and practices that take place daily in every classroom and school shape educational output. Thought of on a district scale it is a complex system in which the processes are not amenable to comprehensive modeling. However, complex systems involve random processes and deterministic elements and models exist to describe the results adequately. These so-called complexity models can allow us to successfully simulate the outcomes with or without complete knowledge of the processes that operate at varying scales throughout the system. It may be small comfort to Eyvind that we offer no advice on key decision points that could have predicted his watery death. However, we could comfort him with a better understanding of chance and determinism that shaped his outcome. A better understanding of the meaning of life does not seem a bad gift to a person at the point of death, the knowledge that gives control over events in his life at this late stage has less value.

This paper is a preliminary exploration into the possibility that a useful mid-level theory in educational reform would be one that models the effect of an iterated pattern of reform adoption on the output that an educational system produces. The choice of the word "useful" is deliberate to express the self-imposed requirement that any mid-level theory should be directed to answering policy questions such as: How many charter schools could be set up in a large urban district before many schools lose all their "good students?" Or, can we find a measure to compare two

school districts' educational performance (with radically different student populations) without resorting to imposing a single set of assessments? The answer using complexity modeling is possibly.

The data from the case studies generated some useful low-level knowledge about reform processes that operate in successful Chicago schools. Our analysis at the school and classroom scales of operation led us to conclude that a simple pattern of reform adoption iterates throughout the educational system at a variety of scales. Additionally, the single-school case data prompted us to think about how that data shared some interesting similarities with issues addressed in the emergent science of complexity modeling. For example, the postponed discussion of site selection that appears after the examination of the data as steady work and local adaptation. This is because that selection process forms an integral part of the data presentation and analysis. Site selection in this paper is an important element in the argument for considering complexity models rather than the traditional preliminary to data presentation.

We hope that using complexity-inspired modeling of educational output we may increase our understanding of the reform process at scales beyond the individual school. We present data and develop an argument that shows that the mid-level process theory Fuhrman and Elmore wanted will remain elusive. But this cross-case data analysis led us to believe we can control reform's output easier than we can control reform's processes, especially on a scale of operation beyond demonstration or experimental sites. The paper concludes with a speculation about what a simulation using complexity theory might look like sketching a procedure to develop such a model. The guiding argument in what follows is that form follows function. When the functions are complex and defy our best efforts to create a model for system wide application there is a need to understand the form. Nobody is denying the importance of studying function, at least as a low-level theory, but studying form represents another way to attack the problem of scaling-up reform efforts and an avenue to develop mid-level theory.

The Illinois Context

This paper uses the findings of the Chicago Panel's case studies conducted for Reform Report that the authors researched in the 1994-95 school year. Illinois accountability measures that were implemented statewide in 1989 significantly influenced Chicago's public schools. However, the city exercised sufficient autonomy to make local adaptations to this state-mandated policy. State legislators singled-out Chicago for special consideration and innovative management experiments. For example, the 1988 Reform Act embarked on what some have called a radical experiment to restructure Chicago schools creating local management councils (Hess, 1991). Consequently, the Chicago Public School system has been no stranger to innovation over the years. We see Chicago's public schools as a "nested" system that receives a series of reform "treatments" that emanate from a variety of scales of operation.

The national scale of operation of the system superficially seems removed from any Chicago classroom's practice and yet we observed in those Mid-west classrooms the federal

system's influence. On any given day identified recipients of various federal initiatives figure in the daily practice of teachers such as Chapter 1 pullout provisions or the presence of special education teachers in "regular" classes. At some of the schools we visited we found instances where the national reform of inclusion had unanticipated consequences on daily teaching practices. In some instances the special education teacher provided an extra resource for all students when integrated into classroom activities. In other instances the inclusion requirement meant that a special education teacher who used to be part of a schoolwide reading program had to withdraw from that activity to make time for testing and counseling duties. Before inclusion these special education teachers in two schools could fit their individual student sessions around a pull-out schedule. Now the school principals required these teachers to be present in regular classrooms when school was in session.

At a state scale of operation another nested sub-system operates. The Illinois State Board of Education created and mandated that schools statewide administer the Illinois Goals Assessment Program (IGAP) in 1989. It is no surprise that we observed instances of the good and bad influence this testing program exerted on teacher's daily practice.

Before 1988 on the district scale, the Chicago district was a highly centralized system that maintained tight control over what happened in Chicago's public schools but with principals exercising creative insubordination on the individual school scale (McPherson & Crowson, 1987). However, the 1988 reforms envisioned the central office as a service center for schools. The district's former superintendents oversaw substantial changes in the central office to make that vision a reality. For example, Superintendent Argie Johnson created several important district-wide initiatives such as building professional networks, highlighting exemplary practices, and emphasizing peer support in failing schools.

The 1988 reforms also established Local School Councils (LSC) at every school. Therefore, each school potentially represents a site for important variation in how they respond to reform challenges. LSCs have three basic governance missions. First, LSCs select and evaluate principals. Second, they approve the state-required School Improvement Plan (SIP). Third, LSCs determine how schools spend their discretionary state chapter one funds, which average about $470,000 per elementary school and $900,000 per high school. Conceptually, the Chicago public system went from a monolithic, hierarchical structure to a more independent unit, flat structure in about six years.

The Chicago Public School system represents 559 (the number of schools in the district) potential study sites within an environment supported or constrained by a single school district and nested in broader systems. The LSC and faculty at each school designed, with varying degrees of success, their unique mix of site-specific inputs incorporating elements originating from other nested sub-systems. Our study focuses on a set of schools whose output can be shown as academic gains in student learning. There were three orientating questions for the individual case analysis; "How has this school achieved academic improvements?" "What measures of success does this school use?" and "What can we take away from this case?" One main question guided

the cross-case study: "What themes cut across individual cases that may lead to mid-level hypotheses about successful reform?"After careful analysis of the low-level processes we concluded that the general theme that showed the most promise for mid-level theory development was to concentrate on the output the site-specific processes produced.

Theoretical Framework

One hope sometimes expressed in the reform community is to aim at systemic reform. This call could be interpreted as a struggle to produce what Rorty (1989) calls the final vocabulary of education. We follow Rorty's view and see education as a field made rich with many vocabularies. Those vocabularies stimulate conversations about reform. Furthermore, the reform vocabularies are constantly under revision as society takes different stances to the three or four (if we include making a profit) basic aims of public education (see Wells 1993). The history of reform, in our framework is a struggle between what is persuasive and therefore desirable in public education that has gone on for around one-hundred and fifty years in many countries.

Tyack and Cuban (in press) and Maclure (1970) in their analyses of the last one hundred years of reform, in America and Britain respectively, support the claim that educational reform is hardly a new idea. Historically theory played an important part in educational reform because it has persuasive power. Reform theory can bring understanding in the sense that a redescription of the phenomena produces what Rorty has described as a "tingle." This tingle comes from the excitement of restructuring our model of the world in a way we judge has greater utility than before. For those who doubt how important this power of ideas is Rorty reminds us, that people and societies die for or dedicate their lives to an idea or insight that produced such a tingle. A cursory glance at the ever-increasing number of "reform case studies" conveys the initial rush of excitement that infects those involved in a particular reform or restructuring effort.

Theory has none of the traditional predictive power according to our Rortian-inspired view. This is fortunate because the ultimate predictive power promised by a final vocabulary would contradict the empirical evidence we constantly read about in the reform literature that local adaptation and people making the reform theirs is what makes "reform work." Also, we could argue that taken as history reform stories show that the "system" has a way of dealing with reforms that for a variety of reasons lose their persuasive power and give way to the next initiative. Some writers would see this as a negative feature of reform efforts symptomatic of the educator's resistance. However, taking a Rortian long-term view as we do to the reform conversation, it is reform's inability to sustain relentless growth that may have some positive benefits over time.

The researchers found that educators in the successful schools reworked any reform initiative into a set of practices that made sense for their school or classroom. The opposite situation would happen if we did realize the goal and produce a final vocabulary. The ability to describe the influencing variables and predict with confidence their outcome, no matter what the specific situations, would silence any argument or contrary view. This is ironic in an endeavor we

have caste as a field made rich by dispute and the clash of ideas.

Guiding Assumptions

Our case evidence agreed with Cohen's (1995) recent observation that there is little direct relationship between policy and practice. Classroom instruction is driven by several related systems such as teacher knowledge or their professional commitments that go beyond any instrument that a policy maker may write at any level in the system. He is not confident about the prospect of the present logic employed in systemic reform ever producing the complete system wide change in instruction the authors of such reforms want.

None of the successful Chicago schools we looked at suggested a set of simple policies or processes that if adopted throughout the district would make every school a desirable learning environment. Kirst's description of state policy creeping into the classroom is a good image of our data. Each teacher we observed or talked to reframed a particular policy and enacted it into their practice according to their sense of what would fit in a particular classroom context. For example, teacher frequently gave brief "policy histories" to make sense of state or district directives. Inevitably, these histories would be highly interpretive, if not factually wrong, in either content or chronology.

Furthermore, teachers cited random events as highly significant. For example, the assignment of a young teacher to a school with only one other teacher close to her age. That random event shaped a successful collaborative relationship between the two third-grade teachers. The story became intriguing when the researchers discovered that two principals had capitalized on this chance event. The principals had created a set of opportunities that strengthened the teaching performance from this collaboration and they monitored its effect on student learning. On the one hand, people interested in promoting reform could point to this as evidence of the benefits of collaboration. On the other hand, a sympathetic critic might ask how we "manipulate random events" and how we insure that collaborative pairs created throughout the district by policy would collaborate (Hargreaves, 1991)? This paper does not provide a new magic variable or a description of the "essence" that distilled into the correct policy implement will produce a systemic reform process. The example captures nicely the need to find a theory that incorporates randomness and determinism.

The researchers chose three persuasive findings, reconfirmed in multiple studies, from the reform implementation literature. These findings guided our analysis and form the requirements for a mid-level model of reform growth across a large number of school sites that analytically form a coherent system[2]. First, as Elmore and McLaughlin (1988) pointed out that the work of carrying out reform is incremental they called it "steady work." It is our contention that the steady

[2]This paper considers mid-level systemic reform at a district level. The authors offer without argument the idea that the family of models we suggest could be applied at a variety of scales. For example, we see it may be possible to compare educational output across several states taking the state as the scale of operation considered.

work is not toward an educational nirvana but a constant historic and contextually-based struggle to use a variety of reform processes. Those processes shape our educational institutions in ways we find persuasive at different times. Second, following the ground breaking Rand's Change Agent study the process of local adaptation works at many levels in the educational system (McLaughlin, 1990). Anyone who has read the reform literature is immediately struck by the power of the local adaptation process, negative and positive, that operates in specific descriptions of reform. Third, Elmore (1979) was the originator of the term "backward mapping." He suggested that policy makers to look at present practices or to consider future practices one wishes to see in classrooms. Then the policy makers write, if it is feasible, policy that supports those practices. Elmore saw backward mapping as a principle for good policy design. However, the idea of backward mapping also reminds us that the outcomes we see operating in the educational system are powerfully shaped by the practices that occur in every classroom on a daily basis. This last thought provides the key insight to understanding our data. Our case data pictures the complex reforms and innovations that shape practice (the process) in Chicago schools thought to be successful on some measure of the school's output. It inspired us to work back from the output to seeing what processes supported that product.

Any model or theory about spreading reform must take into account these three significant features from empirical and theoretical work. In summary our model should depict reform spreading incrementally from one site or person to another and incorporating some notion of contingency or randomness. The model has to adapt to change and not rely on a defined set of variables that produce consistent effects as people adopt a particular reform. Finally, the model has to reflect the practices we see or that we wish to support in teachers' classrooms.

The Shape of Reform in Chicago's Classroom Practices as Steady Work

The analytic task differs from the usual one of the researcher trying to find a single pattern that repeats across cases and building an argument that this pattern is scaled-up to a model that covers a specified range of cases. For example, at a superficial level we could present data that showed that many people in successful schools worked hard to maintain a vision that focused on instructional matters. However, close analysis of the data showed that the visions varied and the methods by which the schools maintained those visions varied across sites.

For example, the instructional vision at one school focused on developing a superior Language Arts curriculum that blended what they felt were the best elements from phonic and literature-based programs. A computer laboratory in this school provided mathematics drill practice that teachers saw as a major contributor to the students' demonstrated learning gains in mathematics. This back-to-basics magnet sharply contrasts with another Chicago school whose vision in language arts had three elements. First, they created a computer laboratory devoted to language arts, especially writing process and ESL (English as a Second Language) activities. Second, they adopted a new literature-based curriculum. Third, the school created an administrative role to meet a new challenge to assimilate a rapidly expanding limited English proficiency student population.

Another school's instructional vision centered on providing their students with a sophisticated level of science experiences that went beyond what the "regular" Chicago elementary student might experience. A core of science teachers built a tradition of innovative teaching in this science and mathematics academy that had partially extended beyond the formal science lessons into other "regular" teacher's classrooms. Two other case schools had significant science programs. However, each of these three schools varied in how they operationalized that vision.

Finally, another school maintained a vision of mastery learning in mathematics and reading that had been diligently followed throughout the school for many years. The teachers used this mastery learning vision to justify a unified set of special reading and writing programs that catered for every ability range represented in the school. One element in this school's program that contributed to its success was the range of special activities that catered for all ability groups including the average-achieving child. The mastery learning vision at this school stood in sharp contrast with another school that used ITBS student gain scores to provide a focus in their conversations to improve teacher practice.

Readers could take away a process here or practice there and implement it in their situation from the individual case studies. However, this provides good evidence for the incremental nature of reform spreading or not spreading between schools. The evidence and the case-level analysis did not justify scaling-up any one of these practices and then applying them in a systematic way throughout the district.

However, these varied examples show the difficulty of scaling-up a model emphasizing a vision with an instructional focus. The patterns we found were significant at each site. But, those descriptions lose significance when stated as a set of general statements about desirable systemic or mid-level theories of procedures and practices. In these Chicago cases we have a pattern of practice that is always incomplete and yet repeats at different scales across the data. The pattern is simple and it consists of complete, incomplete, and nonadoption by teachers of a variety of reform ideas. Any of these case studies could provide "convincing" evidence of a particular reform but the conviction comes at the expense of close examination of the data.

A good example of this simple iterating pattern of reform adoption at different scales is a successful school that followed multiple goals in language arts since the 1988 reforms. The school's improved output was even more remarkable when one considers the school was also responding to an influx of poor children many from Arabic or Hispanic families. The principal felt that the new reading program would help the children at this school achieve the new state language learning goals that she valued. The principal spent considerable organizational resources preparing the teachers for a new reading series that emphasized writing. At a school scale it looks as though we have the example of full implementation. However, the new reading series represented one part of a triple strand strategy aimed at improving all student performance in a situation where student performance might drop. The computer writing laboratory formed the second strand in the school improvement plan. Over a ten-year period, the principal and the

teacher in charge of the program slowly integrated the laboratory's activities into the language program. At the time of data collection the participants quickly acknowledged the process was not complete. Again at a school scale we see a pattern of partial adoption.

A third strand in her improvement plan began many years ago when she became the first non-teaching deputy principal. That role, now occupied by her chosen candidate, was a vital element in coordinating the smooth assimilation of new students who arrived. This structural reform was the easiest element to enact at the school now that the Local School Council could create and fund the post. Again we have a school scale element with full adoption. Several case study schools created a variety of administrative roles with the state Chapter 1 money. Prior events at this school set the precedent for the role (Chicago elementary schools officially have a teaching deputy the school-based councils allocate money to create a non-teaching post). The principal simply rewrote the function in response to recent instability in the student population.

At the individual teacher scale we find another iteration of the simple adoption pattern. A second-grade teacher represented a person who adopted this triple-reform package. We admit the assistant principal is difficult to see absorbed in this teacher's classroom practice. However we have no reason to doubt this external structural change had the effect she claimed. She claimed the new reading and writing materials exposed children in her class to sophisticated content that the students found stimulating. The computer laboratory and the nonteaching deputy all played a significant part in the school's recent success. This second-grade teacher used the new language program and computer lab when we observed her classroom practice.

An eighth-grade teacher confirms his colleagues impression with the improving standards he saw when he compared the students' portfolios. However, he partially adopted the reform package into his practice. He dismissed the computer laboratory's contribution to the improved student writing. The computer program for this teacher simply gave students keyboard experience. He saw no need to incorporate "keyboarding skills" students received in the writing laboratory with what he did in class.

No teacher at the school admitted that they rejected the new language vision outright but several could fall in the category of potential candidates for rejection or active critics. One teacher claimed the new language books had some interesting material in them. However, she had to look to other sources for the material to cover the basic reading skills she valued as part of her practice. Unfortunately, the new series lacked sufficient detail and extra practice. Another skeptic explained that elements of the new language-based approach badly served the needs of poor achieving students. Teacher observations in individual classrooms confirmed a mixed pattern of complete, partial and in some cases rejection of the new language arts practices.

The case showed some slow capacity building among the teachers with some varying success in the computer laboratory and the new literature-based program. A capacity building model lacks general persuasive power in the sense that if one seizes on the notion of capacity building as a mid-level process to spread reform the case data clearly showed mixed outcomes. A

mid-level theory has to accommodate the adoption pattern that occurs at different scales rather than deliberately spread that pattern of inconsistency throughout the system. As David Cohen speculated and this data showed individual teachers draw upon their knowledge, skills, and dispositions to create classroom practices that make sense to them. This case represented a good example of a school engaged in the steady work of reform.

The Shape of Reform in Chicago's Classroom Practices as Local Adaptation

In a second major example from the individual case studies the emphasis changes from the steady work theme to a case of local adaptation. This school began eleven years ago as a back-to-basics academy. Several Chicago schools in the mid 1980's became specialized learning centers for the district. Three individual Panel case studies, faced declining student roles or some other compelling local argument to change from neighborhood elementary schools. These schools (two elementary and one high that occupied a former elementary school) opted to become thematic magnet schools. This decision combined by chance or design with events at the district level. The district, which had to fulfill a recent court agreement, stipulated that these magnet schools drew citywide student applicants from a racially balanced lottery. The magnet school's varied in what they offered their prospective students but with some exceptions these schools supposedly did not select the students they took based on ability.

An interesting feature was the ability of these schools having chosen a theme to select the staff for that program. This ability to choose the initial staff created an opportunity to gather a group of teachers that shared a particular instructional vision. In the two elementary cases the Panel wrote about the science and mathematics academy where the principal retained many original teachers but in the back-to-basics case the principal hired new teachers and he created the staff for the school[3]. A mid-level theory that advocated retaining or replacing teachers to staff a successful innovative program depends in these data on the case chosen and the time at which these data were gathered. We would not endorse that outcome. It clearly shows the dangers of looking toward a single process as the foundation for a mid-level theory. There seems little to guarantee that a policy that allowed schools to choose their staff based on a shared belief system coming in would outperform a school that built a common commitment over time. The ability to hire and fire at a school scale would have to show that the inevitable teacher instability would accrue superior student learning gains.

The back-to-basics academy had a history of strong achievement results in reading and writing. Teachers promised parents that any curriculum change in language arts would be a blend approach would retain the best of phonic and literature-based instruction. Observation of the school's language program revealed varied approaches to phonics and literature-based activities at the classroom scale. A close examination in classrooms revealed a mix of the activities to

---

[3]This school that had, ten years before, taken advantage of the opportunity to hire a staff who shared a vision had to subsequently conform to regular district staffing practices. This small school had lost several of the original key personnel and had to make adjustments to the "special" program they originally designed.

represent this blending idea. Often that classroom scale mixing occurred within a single language instruction period. For example, in a third-grade classroom we observed the teacher started the lesson giving a short vocabulary test emphasizing phonic skills. The teacher then launched into a genre discussion and finished the class with a demonstration of topic sentences for a paragraph exercise she wanted completed for homework.

A sixth-grade teacher interpreted the "new" blended language program a different way. She showed her class a video about the work of the author, Janet Shaw, then she led a class discussion about what they saw. This teacher established an incentive scheme for the children that rewarded them for reading a novel a month. She found that children lacked the information to choose novels and so she gave them examples of good stories to read. In this teacher's case we have a good example of someone integrating a new idea into her practice and that practice being a complex mesh of knowledge, skills and dispositions. The literature-based program influenced her to set up a way to encourage students to read novels. She incorporates a fairly common belief among American teachers that children like incentives or rewards. The back-to-basics tradition at this school laid a heavy emphasis on the teacher's responsibility to be a source of knowledge hence her production of a list of "suitable" or approved literature. At the kindergarten level we found a range of activities that ranged from creative writing and picture activities to "publishing" a phonic rhyming activity as a book.

In this example and all the other schools we visited we found an iterating pattern of local adaptation at the school and the teacher level. This literature-phonic case shows considerable variances among teachers in their adaptive classroom practices. A reasonable question to ask is where is the cutoff point in local adaptation? Specifically in this literature-phonics case where is the boundary at which say a literature-based program no longer looks like what a literature-based advocate would endorse as following the spirit of that particular reform? The single teacher pursuing her advanced degree was the guiding force in the school's move to a literature-based program. It could be fair to characterize this as a chance happening that set a series of events in motion. This second-grade teacher may have come back with some other idea if she had chosen a different thesis topic for her Masters degree. Instead, she returned and persuaded her colleagues and parents about the virtues of a new language program. Except for informal teacher conversations, there was little evidence of a school-wide efforts to produce uniformity of practice on the classroom scale. Instead, a school-scale adoption conversation involved how a literature-based program easily combined with the phonic approach this back-to-basics school already valued. Each teacher adapted the new textbooks and program materials into their classrooms as they saw fit. The examples from classroom observations show the broad variation this adoption strategy produced. This case provided a host of examples of local adaptation at a variety of scales throughout the school.

As a principle of policy design local adaptation's effectiveness on a state scale is the subject of heated debate in Washington, DC and many governor's mansions around the country. Advocates of federal block grants extol the virtues of local adaptation by individual states. As an explanation in research local adaptation has great descriptive power at a variety of scales. Our data provided

plenty of good descriptive examples of local adaptation. So much so that we came away confident that sustained observation in a particular site would yield the same pattern of strong, weak and no implementation in the teacher's practice. Analysis across the successful schools showed this implementation pattern in different lessons by the same teacher, across individual teachers, and school scales of operation. Both chosen cases and others we could describe raise serious doubt about advocating any single reform process model for application as a mid-level theory. There is plenty of important work to develop our understanding of low-level (context-bound) reform. However, our interest lies in producing a mid-level theory or model and to do that we turn our attention to consider mid-level output of what are low-level processes that operate in any educational system.

School Selection and the Measurement Challenge: Applying the Principles of Backward Mapping through output to classroom practices.

Benoit Mandlebrot (Wheatley, 1994) posed a simple challenge to his colleagues and students. How long is the coast of Britain? His colleagues soon realized when they delved into the implications of fractals that there was no final measurement to answer the question. The writers faced a similar challenge as we tried to identify a set of successful schools to study.

Rose (1995) represents the latest in many descriptions of reform and educational success premised on the recommendation strategy. Rose is very careful to point out within the introduction that the descriptions or cases that follow have no easily identified unifying theme. The power and the weakness of the multiple site descriptions are that whatever the author captures at each site is unique to that context. Rose does not reduce what he saw into a call for a particular national reform policy. Instead, Rose produced a multiple-site descriptive case study that has persuasive power in its best sections. However, this book is not a cross-case analysis over multiple sites.

We tried a systematic recommendation strategy, similar to that used by Rose and others, in the Chicago case studies and we generated a list of around forty schools chosen by a variety of people close to Chicago schools. The people making the recommendations represented a broad cross-section drawn from teacher educators responsible for student teacher practice, representatives of reform organizations, local businesses involved in reform efforts, and central office personnel. The significant thing that emerged from this was that despite the large number of schools recommended only one school was mentioned by two independent experts. In addition, a few schools on the recommended list corresponded to those we identified later using the gain score strategy.

The initial reaction is to conclude that the recommendation strategy is a poor "measure" of successful schools. This is a puzzling conclusion when we know that all the people making the recommendations had knowledge of Chicago schools. Sometimes the experts invested money and their professional effort into the schools they identified. If we accept that most recommendations represent some measure of success then that recommendation list resembles an attempt to answer

the Mandelbrot challenge. In the end we rejected using this list to identify elementary schools because typically those recommendations identified schools engaged in interesting reform processes. Our interests lay in finding schools that demonstrated successful student learning.

The recommendation strategy was retained for a list of high schools combined with some weak performance data[4]. In Chicago high schools there are no district-wide test data available that one can feel confident has much relationship to what goes on in those schools. Only one high school became the subject of the individual case studies and the Chicago High School Agricultural Sciences does not qualify as a regular high school on many levels.

In elementary schools, we took the Illinois Goals Assessment Program (IGAP) results and we calculated "gain-scores" for schools. Unfortunately until recently IGAP scores for individuals were not kept. What is the difference between a gain score and between average scores over several years? A gain score comes from a cohort. These cohorts represent students who go through the years together. Essentially, we calculated a gain score by taking the test score at the end of a time and subtracting it from the test score at the beginning.

Unfortunately, this is not a "magic" solution to reporting accurate test scores. In a perfect world, the cohort is completely stable and no students leave or come into the group between the two points. Clearly, this is not so in schools that have many students transferring in and out during the year. We chose schools that had a low student mobility rate to combat this problem. A second problem is that schools can inflate their scores by not including students who might test poorly. We only looked at schools where more than 80% of the eligible students in a grade took the test. A final problem is that many tests change over the years and so scores from different years are not directly comparable. This led to the choice of using the IGAP scores, not the Iowa Test of Basic Skills (ITBS), because IGAP changed little in its form or scoring methods until the 1994-95 school year. The significant difference between these two identification strategies is that the gain-score method identifies in a crude fashion the educational output of nineteen successful elementary schools. The nineteen schools had reading (7) and mathematics (8) gains, with two more schools that appeared on both lists. Other available measures then confirmed the measurable output. For example, we considered writing scores for cohorts where the scores were available. The recommendation method is a measure of identifying forty schools with interesting educational processes. These exercises brought into sharp relief a discontinuity that now exists between interesting educational processes and identifiable output.

Better identification of educational output is at hand. For example, the Chicago Panel has a study undertaken through a subcontract by Bryk and Thum (in press) that uses the Iowa Test of Basic Skills. This work carefully equates ITBS scores from the prereform period to the 1994-95 school year using Rasch measures to account for the different test forms used. The study will document the learning gains of those individuals that remained in one school and the study could

---

[4]Mead has proposed a cluster analysis of high-gain course sequences as a response to this deficit in our knowledge of Chicago high school output.

give interesting information about the learning gains of those that transfer within the system.

An important subsidiary study Kerbow (1995) shows that a significant amount of the student mobility occurs within the Chicago system. Kerbow's study confirms our research strategy that controlled for student mobility when we identified our "successful" schools. Future "output" studies can use the individual student IGAP scores that some feel measure more better that the ITBS the curriculum in place in many Chicago schools. Even further along we might use newly developed performance tests to identify schools that display learning gains.

It is central to our argument that we show this possible progression of measures for successful schools because these measures demonstrate fractal-like properties[5]. There are methods that promise to measure the output of school systems. It is not clear that these output measures can deliver on any claim of final precision. Each version of the measurement instruments gives a different reading and we are forced to consider the possibility, as Mandlebrot's colleagues were, that there is an underlying logic that means the definitive answer to the measurement question will continually elude us.

Where do we go from here?

This cross-case analysis looks as though it reached the same impasse many reform studies reports. However, the findings reported are significant in the sense that our output-driven study confirmed the "solid" knowledge process-driven studies. A key distinction is that these successful schools represented a mix of reform processes and yet they confirmed the outcomes process-driven studies reported. First, we showed that the schools we identified by an available output measure had an instructional focus in their vision. Different "successful" schools chose a variety of instructional visions to form that focus. Furthermore, what an individual teacher enacted in their classroom practice under that visions' rubric varied considerably. Even in successful schools it was possible to see an adoption pattern at different operating scales. The adoption pattern consisted of complete adaptation and adoption into practice, partial adoption but easily extinguished by a new idea, and resistance or rejection of an idea into the teacher's classroom practice.

Second, at least in the schools visited, we confirmed the processes of local adaptation for reform efforts. In the process of local adaptation we saw a fairly simple adoption pattern by individual teachers and within different lessons. As with the steady work of developing and maintaining an educational vision the three-element adoption pattern could be interpreted from what we saw.

---

[5]The particular fractal property is that student learning, like the coastline of Britain, represents an infinite and therefore ultimately nonmeasurable "distance." This does not mean that it is impossible to provide a descriptive quantitative measure of either phenomena. Furthermore, it is not evidence that the search for better methods to measure and identify student learning should be abandoned. Better measurement is a worthy goal it is ultimate measurement that will always elude us.

Third, the measurement of output, although it highlighted some difficult technical challenges yielded the insight that multiple random and determined processes operating in different classrooms produce output that can be rationally argued about, defined and possibly measured. We were struck by the similarity between ourselves with those that try to construct predictive weather models. Our knowledge of the weather processes is always complete and will defy our best technical efforts to predict what will happen. This has led to the serious suggestion that it is easier to control the weather than it is to predict it. This conclusion seems appropriate in the reform context where one hundred and fifty years of knowledge about the process has done little to help our predictive ability except to know how impossible systemic process reform is to achieve.

It is fair to conclude that this research effort in Chicago failed to find the mid-level "process" theory that would allow systemic reform. At an early stage in the individual case study research it was apparent that an output-driven case approach was unlikely to identify a consistent set of practices across cases. The individual case studies provide a collection of promising practices, some that clearly contradict each other, that individual schools or teachers might adapt into their classroom practices. That adoption pattern provides an interesting low-level model of reform where ideas diffuse through the system one school or teacher at a time. However, we set our task as finding a mid-level model for reform.

The researchers, based on the Chicago data, considered nonlinear models used to model complex output. We take backward mapping seriously and so we concluded that the best hope for mid-level theory lies in a model that helps us to understand the form or distribution for the output of reform practices on the ground. This is an important first step to understanding a complex educational system. Later, dependent on the success of this project, it may be possible to simulate growth and decline for learning gain patterns across districts and other large scale areas. All models require a major assumption. Our working assumption was that reform in Chicago represent a mix of deterministic efforts to produce restructuring coupled with the operation of random patterns of adoption. Nonlinear models are emerging objects for study in many fields that we think have application within education. Educational output is a result of the complex processes that go on in schools so we will explore where the thought that form follows function leads.

Instances of Reform as Fractals?

Fractals[6] have two important qualities reflected in the Chicago case study data. The first important fractal quality is that fractals have fractional dimensions. This means that the form they take occupies less than all the available space. The second quality is that we find familiar patterns

[6]Most readers are familiar with the idea of calculating area or volume in two and three dimensions. Areas with fractal or fractional dimensions occupy values greater than one dimension but less that two dimensions. Volumes composed of fractals occupy a space greater than two dimensions but less than the three dimensions traditionally used to describe them. The important point is that a fractal object with fractional dimensions does not occupy all available space.

at different scales. The Chicago reform data had similar qualities. We found that even in supposedly "successful" schools the reform pattern did not occupy all available space. Take for example the school that adopted the "literature-based and phonic blend language program." The researchers easily identified the enthusiastic designer of the school's hybrid literature-based program but as the reform spread throughout other classrooms over a three-year period significant variations occurred. This state of affairs provides poor predictive power for an observer coming into any classroom in that school. What would the observer see? The language lesson data clearly showed the observer could find anything from a recognizable literature reform activity such as a genre discussion going on to a phonic drill. To make matters worse for an observer bent on prediction it was possible that they would see both activities in the same class depending on when the observer sampled the observation.

The guiding assumption in fractal geometry is that form follows function. The theoretical objective is to produce a model that replicates the forms we find in the schools. When we achieve this goal using an appropriate model we can then use that model to simulate the effects of slight changes in the system. A surprising finding in this line of work is that simulating faithfully all the internal variables and processes to recreate a particular form is not necessary or the best way to model that form. In real-world models[7] statistical (using probability to make decisions in the generating phase) fractals play an important part to generate the fractal forms. The realism of a natural scene comes where randomness of form exists within a well-articulated structure. For example, erosion is a complex geographical process that generates the easily recognized fractal form of a dendritic (tree-shaped) drainage basin. A simple equation, squared paper, and dice with four hours work can produce a serious working model for a drainage system that recreates the form but in no way simulates the "real-world" processes that function in a river system. This is a good example combining deterministic processes and random events that combine so that form follows from the functions of the natural system. Reform or innovation can be seen as processes that like erosion combine randomness and determinism that result in an observed form or in the educational sense learning output.

Another striking similarity between education processes and complex systems is that most educational reforms with persuasive power have theoretical underpinning that promise a given structure and a set process that will provide a desirable or observable outcome. In the initial stages of adoption we ignore small perturbations in the use as inevitable short-term problems that eventually get evened out and therefore have no significant long-term effects. In a linear model to describe the spread of reform we attribute the failure of the theory to model the spread of a reform as part of the variance not explained by the multiple elements considered or measurement error. By contrast, in some nonlinear models the operation of chance and small perturbations improves rather than inhibits the model's ability to simulate the form found in the empirical evidence.

---

[7]A real-world model imitates the natural forms found in the data. The fractal models many of us have as screen savers on computers exist in an imagined mathematical space. A better example of a real-world model would be the flight simulator landscapes generated for several popular games.

In the geomorphology and urban geography fields it is possible to go out and collect mountains of data. However, the researcher then finds it hard to visualize how that data should be presented in a way that bears a resemblance to what we understand to be there on the ground. Likewise, education is a field where the data are extensive but they are difficult to visualize as a set of coherent practices occurring uniformly throughout the educational system. However, we should heed the caution that in urban modeling (especially in the 1960-70's) the mathematically elegant models made no sense when geographers compared the "model" with the visual and physical reality. In urban geography this has led to a back-to-basics movement[8] that argues that graphical representation generated by computer should correspond to visual reality rather than statistical reality. The lesson for urban geography parallels the statisticians caution that an elegant model should not be used simply because it is available and mathematically interesting.

A procedure for fractal model building in education

Just as we found it was important that school's had a vision with an instructional focus at each site. Researchers should see this "vision with a policy focus" as a glimpse at the possibilities in fractal modeling[9]. The following represents a speculative exercise from which some important lessons can be drawn. These lessons could help us develop mid-level model for reform output. The model provides a way to gauge how well a system would function under either large-scale application of low-level (school scale) reform theories or application of high-level (state scale) reform models. Simultaneously, to recognize the findings of our cross-case analysis, the fractal and other complexity models account for a mixture of determinism and chance to play a significant role in the pattern of output (learning) that is produced. The mix of determinism and chance reported under the solid findings about the spread of reform, that is the steady work of vision building and the process of local adaptation.

First, we assume that the Chicago Public School district is a system contemplating a new systemic reform initiative that advocates setting up many reform sites to produce systemic change[10]. A major number here represents more than experimental demonstrations in selected areas but not necessarily all the schools in a single system. Second, for the purposes of this example we accept at face value that the proposed reform improves student achievement. We expect that the advocates can offer evidence of learning perhaps from demonstration sites that

---

[8] The authors would be the first to admit the unfortunate coincidence in terminology that a back-to-basics phrase represents in educational circles. The text clearly shows that to an urban geographer the term does not have a negative connotation and the term has a very different nonpejorative history.

[9]Fractal modeling is understood in this context to mean in general terms using a mathematical equation to generate a distribution across a grid that represents a large school district or some other large scale area with many individual units. The purpose of the model is to reproduce the learning gains as a density pattern that corresponds to the patterns observed in reality.

[10]We could also describe a simulation procedure that allowed comparison of district output before and after the imposition of a particular type of instruction. In a Chicago context that could mean before and after direct instruction became district policy.

equate with the system assessments of learning gains available at the time.

Educational reform exhibits a self-similar property, found in objects with fractional dimensions, through its adoption--rejection pattern in teacher practices that we observed in "successful" Chicago schools. Local adaptation produces a fractal pattern of student learning that results from the complex mixing of the three responses in individual teacher practice. The first teachers' response is to incorporate a reform into their practice over a sustained period but recognizing that the classroom practices change or adjust as the teacher incorporates new initiatives. The second teachers' response is to experiment with a new practice but fail to incorporate it in a sustained way in their classroom practice. The third reaction is to reject the reform in their classroom practice, even if they state they are practicing the reform, their rejection shows in the reform's invisibility in classroom practice. This pattern of response at the teacher-individual classroom scale we hypothesize occurs at grade levels, in schools, districts, state and national scales.

The fractional (fractal) pattern of student achievement results from a set of internal and external variables that current research might never completely identify. The confidence that we feel about identifying those variables provides the first important decision for model development between two "families" of models (Batty & Longley, 1994). If we are reasonably confident that we can identify the key variables in the system that influence output then we could create the educational equivalent of the complex weather forecasting model. However if we lack confidence in our ability to identify the relevant key variables we turn to a second model choice that use ergodic measures. Ergodic measure models are useful when we are either unsure of the processes that operate in the system or constructing a model that simulates the variables is not worth the effort involved. It is important to bear in mind that the choice of an ergodic based model is not a "poorer" choice as the experience in weather forecasting teaches us.

Weather forecasting simulations attempt to model climatic and weather conditions by designing vast numbers of small simulation cells with many variables included. Early meteorologists involved in these complex model building attempts hoped for an ever-improving predictive model as we identified all the variables and understood the processes. Recent advances with supercomputers and formidable equations have dramatically increased our short-term predictive ability. However, the work shows that there is approximately a two-week forecasting limit that is unlikely to be broken in the foreseeable future. Later workers acknowledge that pursuing the goal of weather control is probably easier than prediction in this naturally occurring system (Waldrop, 1992). In educational reform the "weather model" lesson is that perfecting control through stronger manipulation of key variables is an appropriate low-level theory goal in selected schools (cells). However, that same weather model starkly highlights the limited predictive power of such a system when we consider 559 very complex cells with a multitude of variables and chance factors operating on how each cell (school and teachers) would react to a particular reform process.

The next step in the output modeling exercise is to map learning gains as two dimensional

"densities." The technical means to do this is at hand using something-like the forthcoming Bryk and Thum equating study for the Iowa Test of Basic Skills. A grid represents each academic year and a square on the grid represents a single school. Each square in the grid is colored according to a scale that represents the percentage of students at that school representing positive learning gains. The result is to have a series of grids (maps) that correspond with the academic years where it is possible to equate the assessment measures[11]. Figure 1 illustrates stylized learning gain grids for academic years' T_0 to T_n

Insert Figure 1 about here

We will assume that after a review of cross-case analysis of reform the ergodic measure model is a better choice than the complex process variable model. In an ergodic model the researchers produce a promising foundational generating model that may not resemble the processes of interest. We gave an earlier example where one of the writers used an ergodic model to reproduce a river drainage pattern "artificially" with no regard to the processes and variables that operated in a "real" river system. Then they choose an appropriate hierarchy. Finally, the researchers introduce random elements in a trial and error procedure to produce the artificial learning gains density landscape.

A good example of this computer mapping strategy is the variants of a Diffusion-Limited Aggregation Model borrowed from physics to successfully produce the patterns of growth or changing densities. Another modeling technique that uses what are called ergodic measures as surrogates for poorly understood processes are the computer generated 3-dimensional landscapes that use a random Serpinski Gasket (simple triangular geometric shape distorted through randomized relaxing of scale, shape and dimension). Batty and Longley note that there is an irony to the use of regular geometric shapes, like triangles in a Serpinski Gasket, to generate irregular morphologies or landscapes with incomplete dimensional qualities (The landscape features do not occupy all available space. Therefore their fractional dimension is greater than two but less than the expected three dimensions).

The significant outcome of that model development is the researcher's ability to calculate what is known as the fractional or fractal dimension (D). Fractal dimensions have a value that in successful models are scale invariant. This D value in a two-dimensional density map has a value between 1 and 2. It makes it possible to measure the degree that the learning gains fill the available space. Obviously, if the model operates it provides a single-value way to compare the "growth" of learning gains over time as the D value increases. Geographers use D value comparisons to express the ruggedness of a coastline or the density of city populations.

---

[11]The exact details for this percent scale are not essential to the point we are making. However, expressing the learning gains as a percentage of the student population making any positive score is justified by the "saw toothing" phenomena Bryk and others have observed in learning gain data. Saw toothing refers to the problem that a large learning gain one year can be followed by a smaller gain in the subsequent measurement periods.

At this point we assume that reformers wish to initiate a major systemic reform effort in the system. The empirical map we created shows schools that have steadily produced learning gains. This map could help potential systemic reformers with a valuable insight. The reform literature shows a depressing tendency for new initiatives to involve the same cast of characters. Bluntly put we reform those that are not in need of reform and we ignore those schools that are unresponsive to reform programs. The series of maps showing school "output" provide a guide to schools that should be chosen for the reform program. Those interested in systemic reform ignore schools where most of their student's showing learning gains and instead concentrate the reform program in those schools that have historically failed to show progress. Our argument here rests on a previous assumption, confirmed by the data, that there is no single set of reform initiatives that if they are uniformly imposed would yield ideal and uniform results in all schools. In other words, concurrent "successful" processes are desirable on a systemic level.

On the model side we could simulate the effects of certain changes in the learning gain patterns. We favor holding reformers to their word and insist they provide output data (learning gains) generated from their model or demonstration sites. Although the simulation does not represent a predictive model, it can display how different reform site placements throughout the district might change the pattern of learning gains.

Batty and Longley report a significant effect of adding determinism to fractal models is that it produces amorphous blobs. Counter-intuitively it is the introduction of random elements to a fractal model that often bring a clear pattern into a sharper focus. In the geographical context amorphous blobs lacking differentiation are not as useful as those same blobs might be in an educational context. The production of blobs would be positive if they show an increasing accumulation of positive learning gains over an area. Again, the fractal dimension calculation can show the effectiveness of the reform program at a variety of scales within the system.

Concerns about simulations that use complexity models

Complexity theory and its modeling strategy is not without its drawbacks or concerns and we would like to mention Some issues require careful study. We would characterize the mathematics and techniques that would allow us to build the fractal models as rapidly evolving and therefore uncertain. There are practical limits for computers to produce a complex simulation with a fine-meshed grid with randomized choice in each square of that grid. Furthermore, the complexity literature presents a lively debate between those who favor very simple equations to model to approximate a complex process and those that favor building complex equations to drive their simulation. Some work on the world's climatic changes through time uses very simple equations with what the researchers hope are key variables not beyond the understanding of a high school student with a good mathematics background (Nicolis & Prigogine, 1989). Other researchers advocate supercomputers and full set of balance equations devoted to modeling faithfully weather and present climatic conditions (Nicolis & Prigogine, 1989). A corresponding complex educational simulation could further our understanding of key educational processes that operate in the system. The question is how much effort and money would be justified by such an

effort?

In contrast to the simulation with complex equations a simple equation strategy raises a substantive concern about ergodic measures that stand in place of complex internal processes that cannot be easily measured. It is frankly unnerving to suggest we use a model that creates the form without due regard to real-world processes. The simulations, at least as far as we have taken it in the paper, are not substitutes for strong educational theory serving as a foundation for a particular reform proposal. Furthermore, the simulation gets strengthened if the model simulates the patterns with better outcome measures. However, on the positive side many potential simulations we could develop rightly emphasize demonstrable output. Recently, critics and advocates of particular reform efforts seem agreed that proposed innovations to be persuasive deliver on an output claim. The claim that X or Y reform represents a promising reform process seems like a recommendation for a trial in a few schools not a persuasive case to advocate full implementation over a wide scale. A fractal simulation that maps patterns of student learning gains clearly satisfies the demand for demonstrable output. Furthermore, we see considerable potential for sophisticated modeling of proposed innovations.

How does a fractal simulation qualify as a mid-level theory?

We see the next task for mid-level theory development is to investigate down the road that this paper simply marks like a signpost to new horizons. A mid-level theory we think must be modest in its goal. The individual case studies clearly showed "successful" schools achieve student learning in many ways. The task for a mid-level theory is to simulate the success of the educational system to deliver what reform throughout history has promised, significant changes in what students learn in school. A mid-level theory of reform cannot tie itself to a single reform program or to a particular method of measuring output.

In our view the search for key reform processes that replicate over multiple sites in reform is a mission bound to failure especially if the goal is to produce consistency in its use. The cases we studied clearly identified processes that were site-specific. For example, once one says that developing a vision with an instructional focus is important the data clearly showed that the vision's content varied across and in the sites. The challenge is that scaling up a site-bound process to higher scales fails to transfer that process in any meaningful way to district or other higher levels. The reform adoption process is not stable across or over time suggesting we are dealing with a dissipative system.

The aim is not to produce a "lazy" model. Those models promised that people faced with the problem use a formulaic response which when applied never worked without local adaptation. Bluntly put some policy makers hoped that a simple formula could serve as a substitute for teachers doing the required hard work to adapt a reform to their local context. The Chicago case data confirmed that to develop and to maintain a vision with an instructional focus teachers engage in a struggle with sustained effort. That effort in successful cases meets, in part, the needs of the particular student population.

This paper started with the idea and then presented some data showing that any mid-level theory or model for reform should adapt to multiple visions with an instructional focus. Second, the model had to fit the process of local adaptation that we saw in our data as a complex mix of randomized and deterministic behavior. Finally, the resulting model if it is useful to policy makers should incorporate the principles in backward mapping. This final point in some ways is the most important in that it reminds us that form follows function. A complex set of seemingly contradictory processes produce the "output" (form) in a large system like Chicago. The models to handle this complexity are at hand. Education as a field of study has always been enriched by looking at other disciplines and fields for ways to resolve difficulties. Geography and physics represent an interesting possibility for further study.

The hundreds of descriptive cases describing reforms teach us valuable lessons that may transfer in their utility to individual schools with suitable adaptation. The accumulating evidence might show the key determinants resist description but there is a modeling strategy using ergodic measures to accommodate that discovery. A mid-level theory should concentrate attention on the form's evolution rather than specifying the specific processes that produced that form. Education reform is a good candidate for the ideas of complexity. If Rorty is persuasive then an obsession with perfecting the reform process is counterproductive to a dialogue with dispute that marks the educational field as special and vital.

References

Batty, M. & Longley, P. (1994). Fractal cities. London: Academic Press, Harcourt-Brace & Co.

Bryk, A. S. & Thum, Y. M. (In progress). Measuring achievement gains in Chicago public schools (tentative). Chicago: Chicago Panel on School Policy.

Ekeland, I. (1993). The broken dice and other mathematical tales of chance (C. Volk, Trans.). Chicago: The University of Chicago Press. (Original work published as Au hasard, 1991)

Elmore, R. F. (1979). Backward mapping: Implementation research and policy decisions. Political Science Quarterly
94 (Winter 1979-80), 601-616.

Elmore, R. F., & Fuhrman, S. H. (1995). Opportunity-to-Learn standards and the state role in education. Teachers College Record, 96, 432-457.

Elmore, R. F., and McLaughlin, M. W. (1988). Steady work: Policy, practice, and the reform of American education. Santa Monica, CA: Rand Corporation.

Hall, G. E. (1995). The local educational change process and policy implementation (pp. 101-123). In D. S. G. Carter & M. H. O'Neill (Eds.), International perspectives on educational reform and policy implementation. London: Falmer Press.

Hargreaves, A. (1985). The micro-macro problem in the sociology of education. In R. Burgess (Ed.), Issues in educational research (pp. 21-47). London: Falmer Press.

Hargreaves, A. (1991). Contrived collegiality: The micropolitics of teacher collaboration (pp. 46-72). In J. Blase (Ed.). The politics of life in schools. Newbury Park: Sage Publications.

Hess, G. A., Jr. (1991) School restructuring, Chicago style. Newberry Park, CA: Corwin Press.

Kerbow, D. (1995). Pervasive student mobility: A moving target for school improvement. Chicago: Chicago Panel on School Policy.

Kirst, M. W. (1995). Overview of recent research on intergovernmental relations. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Maclure, S. (1970). One hundred years of London education 1870-1970. London: Allen Lane, Penguin Press.

McLaughlin, M. W. (1990). The Rand Change Agent study revisited: Macro perspectives and micro realities. Educational Researcher 19(9) 11-16.

McPherson, R. B. & Crowson, R. L. (1987). Sources of constraint and opportunities for discretion in the principalship (pp. 129-156). In J. J. Lane & H. J. Walberg (Eds.). Effective school leadership: Policy and process. Berkeley, CA: McCutchan Publishing Corporation.

Nicolis, G. & Prigogine, I. (1989). Exploring complexity. New York: W. H. Freeman & Company.

Rorty, R. (1989). Contingency, irony, and solidarity. New York: Cambridge University Press.

Rose, M. (1995). Possible lives: The promise of public education in America. Boston: Houghton Mifflin Company.

Tyack, D. & Cuban, L. (In press). Tinkering toward utopia: Reflections on a century of public school reform. Cambridge: Harvard University Press.

Waldrop, M. M. (1992). Complexity : The emerging science at the edge of order and chaos. New York: Touchstone Books.

Wells, A. S. (1993). Time to choose: America at the crossroads of school choice policy. New York: Hill and Wang.

Wheatley, M. J. (1994). Leadership and the new science: Learning about organizations from an orderly universe. San Francisco: Berrett-Koehler Publishers Inc.

Figure 1. Mapping learning gains as densities over time for a "growth" model.